

1 **Evaluating methods to account for system errors in**  
2 **ensemble data assimilation**

3 **Jeffrey S. Whitaker\* and Thomas M. Hamill**

NOAA Earth System Research Laboratory/Physical Sciences Division, Boulder CO

submitted as an article to *Mon. Wea. Rev.*

October 14, 2011

---

\*325 Broadway R/PSD1, Boulder, CO 80305-3328, Jeffrey.S.Whitaker@noaa.gov

Inflation of ensemble perturbations is employed in ensemble Kaman filters to account for unrepresented error sources. We propose a multiplicative inflation algorithm that inflates the posterior ensemble in proportion to the amount that observations reduce the ensemble spread, resulting in more inflation in regions of dense observations. This is justified since the posterior ensemble variance is more affected by sampling errors in these regions. The algorithm is similar to the 'relaxation-to-prior' algorithm proposed by Zhang et al, but it relaxes the posterior ensemble spread back to the prior, instead of the posterior ensemble perturbations.

The new inflation algorithm is compared to the method of Zhang et al, and simple constant covariance inflation using a two-level primitive equation model in an environment that includes model error. The new method performs somewhat better and is less sensitive to variations of the inflation parameter around the optimal value. Combining the new multiplicative inflation algorithm with additive inflation is found to superior to either of the methods used separately.

Tests with large and small ensembles, with and without model error, suggest that multiplicative inflation is better suited to account for un-represented observation network dependent assimilation errors (such as sampling error), while model errors (which do not depend on the observing network) are better treated by additive inflation. A combination of additive and multiplicative inflation can provide a baseline for evaluating more sophisticated stochastic treatments of un-represented background errors. This is demonstrated by comparing the performance of a stochastic kinetic energy backscatter scheme with additive inflation as a parameterization of model error.

# 1. Introduction

The ensemble Kaman filter (EnKF), an approximation to the Kaman filter, estimates the background-error covariance from an ensemble of short-term model forecasts. The use of EnKF data assimilation systems to initialize ensemble weather predictions is growing (e.g. Whitaker et al. 2008; Hamill et al. 2011; Houtekamer et al. 2005, 2009; Buehner et al. 2010), because of the simplicity of the algorithm and its ability to provide *flow-dependent* estimates of background and analysis error. In the EnKF, it is assumed that the background (prior) ensemble samples all sources of error in the forecast environment, including sampling error due to limitations in ensemble size, and errors in the model itself. Inevitably, some sources of error will be under-sampled, resulting in a EnKF with a suboptimal estimate of the background-error covariance with systematically underestimated variances. Such an EnKF may not give enough weight to observations, which in a chaotic system will cause the subsequent ensemble forecasts to drift farther from the truth. At the next assimilation time, ensemble-estimated covariance model will be even more deficient, causing the update to give even less weight to observations. This problem can progressively worsen, potentially resulting in a condition called “filter divergence”, in which the ensemble variance becomes vanishingly small and observation information is completely ignored. Because of this, all EnKF systems used in weather prediction employ methods to account of unrepresented or underestimated error sources in the prior ensemble. These include multiplicative inflation (Anderson and Anderson 1999), which inflates either the prior or posterior ensemble by artificially increasing the amplitude of deviations from the ensemble mean, and additive inflation, which involves adding random perturbations with zero mean from a specified distribution to each ensemble mem-

ber (Mitchell and Houtekamer 2000). Whitaker et al. (2008) compared simple uniform multiplicative inflation with additive inflation in a simple model, and found that additive inflation performed better, since the simple uniform multiplicative inflation generated too much spread in regions less constrained by observations. Houtekamer et al. (2009) compared additive inflation with various methods for treating model error within the forecast model itself, such as multi-model ensembles, stochastic-backscatter (Shutts 2005; Berner et al. 2009) and stochastically perturbed physics tendencies (Buizza and Palmer 1999). They found that additive inflation, sampling from a simple isotropic covariance model, had the largest positive impact. However, Hamill and Whitaker (2010) found that parameterizing unrepresented error sources with additive inflation will decrease the flow-dependence of background-error covariance estimates and reduce the growth rate of ensemble perturbations, with potentially negative consequences on analysis quality.

In this study, we re-examine the use of inflation (additive and multiplicative) as methods for accounting for under-represented sources of background-error in ensemble data assimilation. The goal is to elucidate the strengths and weakness of each method in isolation, and justify the use of both simultaneously. Experiments are conducted using an idealized 2-level primitive equation model on a sphere, including model error. We hypothesize that multiplicative inflation algorithms should inflate more where observations are dense to account for the fact that sampling errors (and other sources of under-represented observation network dependent data assimilation errors) are likely to be a larger fraction of the total background error in those regions. To this end, we propose a very simple algorithm that inflates the posterior ensemble proportional to the amount that ensemble variance is reduced by the assimilation of observations, and compare this new algorithm to

existing ones. We also hypothesize that additive inflation will outperform multiplicative inflation alone when unrepresented model errors dominate unrepresented observation-network dependent system errors (which in this simplified environment consists solely of sampling error due to limitations in ensemble size). The opposite should be true when sampling error dominates model error. When neither model error or sampling error dominates, a combination of multiplicative inflation and additive inflation should perform better than either alone. To put it simply, we aim to demonstrate that when using inflation, observation network dependent assimilation errors are best handled by multiplicative schemes, while model errors (which do not depend on the observing network) are best treated by additive schemes. The following section describes the algorithms used and experiments performed, while the results and conclusions are summarized in the final section.

## 2. Idealized experiments

### *a. Forecast model*

The forecast model used in these experiments is virtually identical to the two-level primitive equation spectral model of Lee and Held (1993). This model was also used in the data assimilation experiments of Whitaker and Hamill (2002) and Hamill and Whitaker (2010). Here, unless otherwise noted, data assimilation experiments are run with a spectral resolution of T31 (triangular truncation at total wavenumber 31), with the two levels set to 250 and 750 hPa. Observations are sampled from a nature run using the same model, but at T42 resolution. The prognostic variables of the forecast model are baroclinic and barotropic vorticity, baroclinic divergence, and barotropic

potential temperature. Barotropic divergence is identically zero, and baroclinic potential temperature (static stability) is kept constant at 10K. Lower-level winds are mechanically damped with an e-folding timescale of 4 days, and barotropic potential temperature is relaxed back to a radiative equilibrium state with a pole-to-equator temperature difference of 80K with a timescale of 20 days. The radiative equilibrium profile of Lee and Held (1993) (equation 3) was used.  $\nabla^8$  diffusion was applied to all the prognostic variables, the smallest resolvable scale is damped with an e-folding timescale of 3 hours (6 hours for the nature run). Time integration is performed with a 4th-order Runge-Kutta scheme with 18 time steps per day at T31 resolution, and 30 per day at T42 resolution. The error doubling time of the T31 model is approximately 2.4 days. The climate of the model (computed as a zonal and time mean over 1000 days of integration) is shown in Fig. 1 for the T31 forecast model and the T42 nature run. The time-mean systematic error of the T31 model is quite small outside the tropics and polar regions.

#### *b. Data assimilation methodology*

The serial ensemble square-root filter algorithm of Whitaker and Hamill (2002) is used in conjunction with a 20-member ensemble, unless otherwise noted. Details are provided in Hamill and Whitaker (2010). Covariance localization (Hamill et al. 2001) is used to ameliorate the effects of sampling error, using the compact Gaussian-like polynomial function of Gaspari and Cohn (1999). Unless otherwise noted, the covariance localization was set so that increments taper to zero 3500 km away from observation locations. This is close to the optimal value for all of the experiments with 20 member ensembles. Observations of geopotential height at 250 and 750 hPa are assim-

ilated at Northern Hemisphere radiosonde locations (Fig. 2) every 12 hours with an observation error standard deviation of 10 meters. The observing network is made hemispherically symmetric by reflecting the Northern Hemisphere radiosonde locations into the Southern Hemisphere, resulting in a network with 1022 observing locations.

*c. Comparison of multiplicative inflation methods.*

Sacher and Bartello (2008) showed that sampling error in the estimate of the Kaman gain should be proportional to the amplitude of the Kaman gain itself, so that more inflation is needed when observations are making large corrections to the background. Therefore, it seems desirable to have a multiplicative inflation scheme that inflates the ensemble variance more in regions where observations have a larger impact.

Zhang et al. (2004) proposed an alternative to simple constant covariance inflation that relaxes posterior (analysis) perturbations back toward the prior (first guess) perturbations independently at each analysis point via

$$\mathbf{x}_i'^a \leftarrow (1 - \alpha)\mathbf{x}_i'^a + \alpha\mathbf{x}_i'^b, \quad (1)$$

where  $\mathbf{x}_i'^a$  is the deviation from the posterior ensemble mean for the  $i^{th}$  ensemble member, and  $\mathbf{x}_i'^b$  is the deviation from the prior ensemble mean for the  $i^{th}$  ensemble member. We refer to this method as “relaxation-to-prior perturbations” (RTPP). Unlike constant covariance inflation, this technique has the desired property of increasing the posterior ensemble variance in proportion to the amount that the assimilation of observations has reduced the prior variance. In the limit that  $\alpha$  approaches 1.0, the posterior ensemble is completely replaced the prior ensemble. For values of  $\alpha$

124 between 0 and 1, part of the posterior ensemble is replaced by the prior ensemble. This approach  
 125 amounts to a combination of multiplicative inflation (in which the inflation factor is less than 1)  
 126 and additive inflation where the perturbations are taken from the prior ensemble. Here we propose  
 127 a new approach, which we call “relaxation-to-prior spread” (RTPS), that is a purely multiplicative  
 128 inflation. Instead of relaxing the posterior *perturbations* back to their prior values at each grid  
 129 point as in RTPP, we relax the ensemble *standard deviation* back to the prior via

$$\sigma^a \leftarrow (1 - \alpha)\sigma^a + \alpha\sigma^b, \quad (2)$$

130 where  $\sigma^b \equiv \sqrt{\frac{1}{n-1} \sum_{i=1}^n \mathbf{x}_i'^{b2}}$  and  $\sigma^a \equiv \sqrt{\frac{1}{n-1} \sum_{i=1}^n \mathbf{x}_i'^{a2}}$  are the prior and posterior ensemble standard  
 131 deviation (spread) at each analysis grid point, and  $n$  is the ensemble size. This formula can be  
 132 rewritten

$$\mathbf{x}_i'^a \leftarrow \mathbf{x}_i'^a \left( \alpha \frac{\sigma^b - \sigma^a}{\sigma^a} + 1 \right). \quad (3)$$

133  
 134 For a given value of  $\alpha$ , the multiplicative inflation is proportional to the amount the ensemble  
 135 spread is reduced by the assimilation of observations, normalized by the posterior ensemble spread.  
 136 Anderson (2009) proposed a Bayesian algorithm for estimating a spatially and temporally varying  
 137 field of covariance inflation as part of the state update. When run as part of an EnKF assimilation  
 138 system using a global general circulation model with all “conventional” (i.e. non satellite radiance)  
 139 observations, the Bayesian algorithm produces a spatial field of inflation that looks very similar to  
 140 that implied by RTPS inflation (equation 3), with large values of inflation in regions of dense and/or

accurate observations, like North America and Europe (Fig. 13 in Anderson et al. (2009)).

The role of covariance localization is to ameliorate the effects of sampling error, yet we have hypothesized that spatially varying covariance inflation is also necessary to deal with the observation-network dependent effects of sampling error. Why do we need both? Covariance localization, as it is typically formulated, tapers increments with distance from the observation localization, allowing the full increment to be applied at the observation location and no increment past a specified cutoff distance from the observation. Allowing the full increment to be applied at the observation location (i.e. having the localization function peak at unity) implicitly assumes sampling errors in the estimation of background-error covariances between model and observation priors that are co-located in space are zero. In the simple case where observation operator is the identity matrix (model state variables are observed), this implies that covariance localization only deals with sampling errors in the estimation of background-error covariances, not variances. Covariance inflation is therefore needed to account for sampling error in the estimation of background-error variances, which will invariably be underestimated by small ensembles. In the case of RTPP inflation, it can be demonstrated that this is equivalent to applying a covariance localization function that peaks at a value of  $1 - \alpha$  at the observation location when updating ensemble perturbations (see Appendix A for details).

Experiments were conducted with three different methods of multiplicative inflation (simple covariance inflation, RTPP and RTPS) to account for background errors not accounted for by the first-guess ensemble, which in this case includes both sampling error and model error, since the ensemble size is small relative to the dimension of the forecast model, and the forecast model is

run at lower resolution than the model used to generate the observations. Ensemble mean error and spread are calculated using the total energy norm

$$E(\vec{V}_1, \vec{V}_2, \theta_{3/2}) = \frac{1}{2}(u_1^2 + v_1^2) + \frac{1}{2}(u_2^2 + v_2^2) + \frac{\Delta\bar{\pi}}{\Delta\bar{\theta}}\theta_{3/2}^2, \quad (4)$$

where  $\Delta\bar{\theta}$  is the constant static stability (10K),  $\Delta\bar{\pi}$  is the difference in Exner function between the lower level (750 hPa) and the upper level (250 hPa),  $\vec{V}_1 = (u_1, v_1)$  is the lower level horizontal velocity vector, and  $\vec{V}_2 = (u_2, v_2)$  is the upper level velocity, and  $\theta_{3/2}$  is the barotropic, or mid-level potential temperature. Ensemble mean error is computed by replacing the velocity and potential temperature in equation 4 by the difference between the ensemble mean and the truth (as defined by the T42 nature run). The ensemble spread is computed by replacing the velocity and potential temperature in equation 4 by the difference between each ensemble member and the ensemble mean, then summing over each ensemble member and dividing by the number of ensemble members minus one. Global and time means of the resulting quantities are computed, and a square root is then applied so that the result has units of meters per second.

Fig. 3 shows ensemble mean background error and spread statistics collected over 1000 assimilation times for the three experiments, after a spinup period of 50 days. The RTPS inflation method produces more accurate analyses and short-term forecasts than either RTPP and constant covariance inflation and is less sensitive to variations of the inflation parameter about its optimal value (the value at which the ensemble mean error is minimized). RTPP outperforms constant covariance inflation but produces very large errors when the inflation parameter exceeds the optimal value. The ensemble spread for all three experiments is less than the ensemble mean error when

ensemble mean error is at its minimum (the dashed lines are below the solid lines at the minimum in the solid line), indicating the ensembles are slightly under-dispersive when they are optimized for ensemble mean error. Both the ensemble spread and ensemble mean error for the RTPS inflation appears to be less sensitive to variations in the inflation parameter. For reference, we also show in Fig. 3 the ensemble mean error and spread for an experiment using the adaptive inflation algorithm of Anderson (2009) (the horizontal cyan curves). The adaptive inflation algorithm requires very little tuning (there is some sensitivity to the value of the prior inflation variance chosen), and produces analysis of similar quality to the best-tuned RTPS results.

RTPP inflation has at least one desirable property; it produces ensemble perturbations that grow faster than the other inflation methods. This is illustrated in Fig. 4, which shows that ratio of background spread to analysis spread for the experiments depicted in Fig. 3. At the minimum in ensemble mean error, the RTPP ensemble spread grows about 19% during over the assimilation interval (12 hours), compared to 7.6% for RTPS inflation and 6.5% for simple covariance inflation. The reason for this can be understood by noting that RTPP inflation involves adding scaled prior perturbations to the posterior ensemble. When the inflation parameter  $\alpha$  is 1, the posterior ensemble is completely replaced by the prior ensemble. In that case, the structure and amplitude of the ensemble perturbations is not modified during the assimilation and the perturbations are simply re-centered around the updated ensemble mean. The assimilation cycle then becomes similar to the process used to compute the leading Lyapunov vector (Legras and Vautard 1995), which reflect the dominant instabilities of a dynamical system. This also explains why the performance of RTPP inflation degrades rapidly when the inflation parameter is increased above the optimal

value - the ensemble perturbations become increasingly co-linear as they collapsing to the leading Lyapunov vector, reducing the effective number of degrees of freedom spanned by the ensemble. However, Fig. 4 shows that the spread growth does not increase for RTPP inflation monotonically as the inflation parameter is increased. This is because the amplitude of the ensemble perturbations becomes large enough that nonlinear effects begin to cause saturation.

To further explore the impact of the multiplicative inflation method on the growth properties of the analysis ensemble, we have calculated the analysis-error covariance singular vector (AECSV) spectrum following the methodology of Hamill et al. (2003). The AECSVs are the structures that explain the greatest forecast variance and whose initial size is consistent with the flow-dependent analysis-error covariance statistics of the data assimilation system. Figure 5 confirms that the RTPP ensemble AECSV spectrum is steeper, with more of the variance concentrated in fewer, faster growing modes. The leading AECSV for the RTPS ensemble grows just as rapidly as the leading AECSV in the RTPP ensemble, but the trailing ones grow more slowly. This results in less spread growth over the assimilation interval, but an ensemble that can effectively span a larger portion of the space of possible analysis errors.

#### *d. Combined additive and multiplicative inflation.*

In Hamill and Whitaker (2005), it was found that additive inflation performed better than constant covariance inflation in a idealized 2-layer primitive equation model, including truncation model error. Similarly, Whitaker et al. (2008) found that additive inflation outperformed constant covariance inflation and RTPP inflation in a full global numerical weather prediction system.

222 Given that RTPS inflation generally performs better than RTPP and constant covariance inflation,  
 223 how does it perform compared to additive inflation? Here we use random samples from a climato-  
 224 logical distribution of actual 12-h forecast model error for our additive inflation. The distribution  
 225 is computed using the same method as Hamill and Whitaker (2005), that is by truncating the T42  
 226 nature run to T31, running 12-h forecasts at T31 and computing the difference between these fore-  
 227 casts and the corresponding T31 truncated nature run fields. The only source of error in these  
 228 forecasts is due to the lower resolution of the forecast model. At each analysis time, 20 samples  
 229 are chosen randomly from this distribution, the mean is removed, and the resulting fields are scaled  
 230 and added to each ensemble member. Figure 6 shows the ensemble background error for experi-  
 231 ments using a combination of this additive inflation and RTPS multiplicative inflation. The additive  
 232 inflation parameter is simply the scaling factor applied to the randomly chosen truncation model  
 233 error fields. The values of ensemble mean error when the additive inflation parameter is zero are  
 234 identical to those shown in Fig. 3 (the solid red line). From this plot, it is easy to see that additive  
 235 inflation without multiplicative inflation produces lower errors than multiplicative inflation alone,  
 236 in agreement with the results of Hamill and Whitaker (2005) and Whitaker et al. (2008). However,  
 237 a combination of additive and multiplicative inflation produces lower errors than either method  
 238 used alone. The minimum error ( $8.6 \text{ ms}^{-1}$ ) occurs with a multiplicative inflation parameter of 0.5  
 239 and an additive inflation parameter of 1.4. Conditioning the additive perturbations to the dynamics  
 240 by adding them to the previous ensemble mean analysis (instead of the current analysis) and evol-  
 241 ving them forward in time one assimilation interval (as suggested by Hamill and Whitaker (2010))  
 242 reduces the minimum error slightly, by approximately 2-3% (not shown). Using random samples

of 12-h differences drawn from a T31 model run works nearly as well as using actual truncation model error fields for the additive inflation, yielding a minimum error of ( $8.8 \text{ ms}^{-1}$ ) when the additive inflation parameter is 0.24 and the multiplicative inflation parameter is 0.5 (Fig. 7).

The fact that a combination of additive and multiplicative inflation works better than either alone suggests that they are representing different unrepresented background-error sources. RTPS multiplicative inflation is by design dependent on the observation network, while the additive inflation we have used is independent of the assimilation system. Therefore, we hypothesize that RTPS multiplicative inflation is useful in capturing the especially deleterious effects of sampling errors in regions where observations are dense, while additive inflation is useful in capturing sources of background error that are assimilation-system independent, such as errors in the forecast model.

To test this idea we ran two experiments, one in which the model error was eliminated by using the T42 model in the assimilation, and another in which the sampling error was reduced by increasing the ensemble size from 20 to 200. In the former experiment, we expect that the relative impact of additive inflation would be reduced relative to multiplicative inflation, since the only source of unrepresented source of error (sampling error) comes from the data assimilation system itself. In the latter experiment, sampling error is greatly reduced, so that the dominant unrepresented source of error should be model error and the impact of multiplicative inflation should be reduced relative to additive inflation. These expectations are confirmed in Figs. 8 and 9. Figure 8 shows that in the absence of model error, multiplicative inflation alone outperforms any combination of multiplicative and additive inflation. Figure 9 shows that when model error is the dominant source of unrepresented background errors, additive inflation alone outperforms any

combination of multiplicative and additive inflation.

*e. Replacing additive inflation with stochastic backscatter.*

The additive inflation algorithm used here is somewhat ad-hoc, and it would be preferable to incorporate a physically based parameterization of model error directly into the forecast model. Such a parameterization would account for the presence of model error directly in the background ensemble forecast. The only source of error in our two-level model experiments is associated with model truncation. More specifically, model error in our experiments is a result of the effects of unresolved and unrealistically damped scales on the resolved scales through an inverse energy cascade. This is exactly the sort of model error that stochastic kinetic energy backscatter (SKEB) schemes (Shutts 2005; Berner et al. 2009) were designed to represent. The algorithm described by Berner et al. (2009) involves generating a random streamfunction pattern from an AR-1 process with a specified timescale and covariance structure. These random patterns are then modulated by the model's kinetic energy dissipation rate (resulting from the  $\nabla^8$  hyperdiffusion). The resulting tendencies are added as forcing term in the vorticity equation. Figure 10 shows the the total kinetic energy spectra for the T42 model, the T31 model without SKEB, and the T31 model with SKEB. The kinetic energy in the T31 model without SKEB is deficient relative to the T42 model at all scales, but especially so near the truncation wavenumber where the hyperdiffusion is active. Adding SKEB to the T31 model brings the energy up much closer to the level of the T42 model. The random streamfunction pattern used to generate the SKEB forcing was assumed to be spatially white in the streamfunction norm, with a decay timescale of 6 hours. The amplitude of the random

streamfunction pattern was set to 15, a value chosen to give the best fit to the T42 model kinetic energy spectrum shown in Fig. 10.

Figure 11 show the results for a set of assimilation experiments using a combination of SKEB to represent model error, and multiplicative inflation to represent other sources of unrepresented background errors (in this case, primarily sampling errors). Not surprisingly, a combination of SKEB and multiplicative inflation turns out to be better than either alone. However, comparing Fig. 11 to Fig. 7, SKEB does not seem to perform significantly better than simple, ad-hoc additive inflation. Also, in contrast to the additive inflation case, SKEB alone does not perform better than additive inflation alone. Of course, there are several tunable parameters in the SKEB scheme (including the total variance injected, the time-scale of the random streamfunction pattern, and the covariance structure of the random streamfunction pattern) and it likely that better results could be obtained by more carefully tuning these parameters. However, our results do suggest that it is surprisingly hard to beat a combination of simple additive and multiplicative inflation as a parameterization of unrepresented sources of error in an ensemble data assimilation system.

### 3. Conclusions

In the EnKF, it is assumed that the background (prior) ensemble samples all sources of error in the forecast environment, including those associated with the data assimilation itself (such as sampling error due to finite ensemble size, mis-specification of observation errors and errors in forward operators) as well as errors in the forecast model itself. We have proposed a new multiplicative

inflation algorithm to deal with the effects of these unrepresented sources of error that is simple to implement in complicated models. Using idealized experiments with a two-level spherical primitive equation model, where the only source of model error is associated with model truncation, and the only source of data assimilation error is associated with finite ensemble size, we show that this new inflation scheme performs as well or better than other commonly used schemes. It has the desirable property of inflating more strongly where the assimilation of observations has a larger effect on the ensemble variance. It is in these regions where sampling error is expected to be a larger fraction of the total background error.

Combining this new multiplicative inflation algorithm with additive inflation, it is found that a combination of the two performs better than either alone, even when the additive perturbations are drawn from an ad-hoc distribution that does not directly use knowledge of the known properties of the model error in this simplified environment. This leads us to hypothesize that multiplicative inflation is best suited to account for unrepresented observation network dependent assimilation errors (in this case sampling error), while model errors (which do not depend on the observing network) are best treated by additive inflation, or stochastically within the forecast model itself. Since the additive inflation algorithm is somewhat ad-hoc, it is expected that a more physically based parameterization of model error, such as stochastic kinetic energy backscatter, will perform better. Tests replacing additive inflation with SKEB in the data assimilation show that it is surprisingly hard to improve upon additive inflation. This suggest that a combination of simple ad-hoc additive inflation with the new multiplication inflation algorithm proposed here can provide a rigorous baseline for testing new more sophisticated representations of unrepresented sources of error in

ensemble data assimilation systems.

More generally, these results suggest that it is desirable to treat different sources of un-represented of background error in ensemble data assimilation systems separately, using as much a-priori knowledge regarding the characteristics of these errors as possible. In the case of inflation, using the fact that we expect part of the missing error to be observation network dependent and part of it to be independent of the observing network leads us to an improved scheme that has both additive and multiplicative aspects. Applying this philosophy to the model error, we might expect that errors associated with convection, boundary layer physics and unresolved dynamics might best be treated separately, as long we have a prior knowledge about the characteristics of these separate sources of error to guide us. Similarly, for unrepresented sources of error associated with the data assimilation system itself, such as mis-specification of observation errors and errors in forward operators, there may be methods that work better than the RTPS multiplicative inflation used here. More research is certainly needed to understand what the most important un-represented sources of error are in operational ensemble data assimilation systems, and how to characterize those errors individually.

## Appendix: Equivalence between covariance localization and RTPP

### inflation.

In the serial ensemble square-root filter (Whitaker and Hamill 2002) the ensemble-perturbation update is given by

$$\mathbf{x}'^a = \mathbf{x}'^b - \tilde{\mathbf{K}}\mathbf{H}\mathbf{x}'^b, \quad (5)$$

where

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} \quad (6)$$

is the Kalman gain,  $\mathbf{P}^b \mathbf{H}^T = \overline{\mathbf{x}'^b (\mathbf{H} \mathbf{x}'^b)^T} \equiv \frac{1}{n-1} \sum_{i=1}^n \mathbf{x}'_i{}^b (\mathbf{H} \mathbf{x}'_i{}^b)^T$ ,  $\mathbf{H} \mathbf{P}^b \mathbf{H}^T = \overline{\mathbf{H} \mathbf{x}'^b (\mathbf{H} \mathbf{x}'^b)^T} \equiv \frac{1}{n-1} \sum_{i=1}^n \mathbf{H} \mathbf{x}'_i{}^b (\mathbf{H} \mathbf{x}'_i{}^b)^T$ ,  $n$  is the ensemble size and  $\tilde{\mathbf{K}}$  is the gain used to update deviations from the ensemble mean. Note that an over-bar used in a covariance estimate implies a factor of  $n - 1$  instead of  $n$  in the denominator, so that the estimate is unbiased. If  $\mathbf{R}$  is diagonal, observations may be assimilated serially, one at a time, so that the analysis after assimilation of the  $N$ th observation becomes the background estimate for assimilating the  $(N + 1)$ th observation (Gelb et al. 1974). With this simplification,  $\tilde{\mathbf{K}}$  may be written as

$$\tilde{\mathbf{K}} = \left( 1 + \sqrt{\frac{R}{\mathbf{H} \mathbf{P}^b \mathbf{H}^T + R}} \right)^{-1} \mathbf{K}, \quad (7)$$

for an individual observation, where  $R$  and  $\mathbf{H} \mathbf{P}^b \mathbf{H}^T$  are scalars, while  $\mathbf{K}$  and  $\tilde{\mathbf{K}}$  are vectors of the same dimension as the model state vector. Covariance localization is applied by tapering  $\mathbf{K}$  to zero

353 with distance from the observation, so that Eq. 5 can be written as

$$\mathbf{x}'^a = \mathbf{x}'^b - \left(1 + \sqrt{\frac{R}{\mathbf{H}\mathbf{P}^b\mathbf{H}^T + R}}\right)^{-1} \Gamma \mathbf{K} \mathbf{H} \mathbf{x}'^b, \quad (8)$$

354 where  $\Gamma$  is a vector containing the localization function on the model grid for an individual ob-  
 355 servation.  $\Gamma$  is unity at the observation location, and zero beyond a specified distance from the  
 356 observation. In the special case where only a single observation is assimilated, RTPP inflation is  
 357 equivalent to multiplying the second term on the right-hand side of Eq. 8 by  $1 - \alpha$ , so that when  
 358  $\alpha = 1$ , the analysis perturbation is identical to the background perturbation, and when  $\alpha = 0$ , no  
 359 inflation is applied. If the factor  $1 - \alpha$  is incorporated into the localization function  $\Gamma$ , it becomes  
 360 apparent that RTPP inflation is equivalent to applying covariance localization with a localization  
 361 function that peaks at  $1 - \alpha$  at the observation localization in the ensemble perturbation update  
 362 (but not in the ensemble mean update). When more than one observation is assimilated, applying  
 363 RTPP inflation after the analysis step is not formally identical to localizing with a function that  
 364 peaks at  $1 - \alpha$  during the serial assimilation update. However, numerical experimentation shows  
 365 that results shown in Fig. 3 are essentially unchanged if RTPP is applied in this way. We note that  
 366 RTPS inflation is not equivalent to a modified localization in the perturbation update.

## References

- Anderson, J. L., 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus A*, **61**, 72–83.
- Anderson, J. L. and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758.
- Anderson, J. L., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The data assimilation research testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296.
- Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ecmwf ensemble prediction system. *J. Atmos. Sci.*, **66**, 603–626.
- Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: Intercomparison of variational data assimilation and the ensemble kalman filter for global deterministic nwp. part i: Description and single-observation experiments. *Mon. Wea. Rev.*, **138**, 1902–1921.
- Buizza, R. and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.
- Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
- Gelb, A., J. F. Kasper, R. A. Nash, C. F. Price, and A. A. Sutherland, 1974: *Applied Optimal Estimation*. M. I. T. Press, 374 pp.

386 Hamill, T., C. Snyder, and J. S. Whitaker, 2003: Ensemble forecasts and the properties of flow-  
387 dependent analysis-error covariance singular vectors. *Mon. Wea. Rev.*, **131**, 1741–1758.

388 Hamill, T. M. and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in  
389 ensemble data assimilation: A comparison of different approaches. *Mon. Wea. Rev.*, **133**, 3132–  
390 3147.

391 Hamill, T. M. and J. S. Whitaker, 2010: What constrains spread growth in forecasts initialized  
392 from ensemble Kalman filters? *Monthly Weather Review*, **138**, to appear.

393 Hamill, T. M., J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011: Global ensemble predictions  
394 of 2009’s tropical cyclones initialized with a ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, to  
395 appear.

396 Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background  
397 error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790.

398 Houtekamer, P. L., H. L. Mitchell, and X. Deng, 2009: Model error representation in an operational  
399 ensemble kalman filter. *Mon. Wea. Rev.*, **137**, 2126–2143.

400 Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen,  
401 2005: Atmospheric data assimilation with an Ensemble Kalman Filter: Results with real obser-  
402 vations. *Mon. Wea. Rev.*, **133**, 604–620.

403 Lee, S. and I. M. Held, 1993: Baroclinic wave packets in models and observations. *J. Atmos. Sci.*,  
404 **50**, 1413–1428.

- 405 Legras, B. and R. Vautard, 1995: A guide to Lyapunov vectors. *Proc. ECMWF Seminar on Pre-*  
406 *dictability*, Reading, United Kingdom, ECMWF, Vol. 1, 143–156.
- 407 Mitchell, H. L. and P. L. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Mon. Wea. Rev.*,  
408 **128**, 416–433.
- 409 Sacher, W. and P. Bartello, 2008: Sampling errors in ensemble kalman filtering. part i: Theory.  
410 *Mon. Wea. Rev.*, **136**, 3035–3049.
- 411 Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems.  
412 *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102.
- 413 Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observa-  
414 tions. *Mon. Wea. Rev.*, **130**, 1913–1924.
- 415 Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation  
416 with the ncep global forecast system. *Mon. Wea. Rev.*, **136**, 463–482.
- 417 Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability  
418 on convective-scale data assimilation with an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **132**,  
419 1238–1253.

## List of Figures

- 1 Zonal and time mean zonal wind (left) and potential temperature (right) for the two-level model run at T31 and T42 resolution. 26
- 2 Observation locations (black dots) superimposed upon a snapshot of barotropic potential temperature (shaded contours every 5K) and upper-level winds (vectors) from the T42 nature run used in the two-level model data assimilation experiments. Continental outlines are shown for reference, even though the model has no orography or land-sea contrast. Data is plotted on a azimuthal equidistant map projection centered on the North Pole. 27
- 3 Ensemble mean background error and spread (in terms of the square-root of total energy norm in  $ms^{-1}$ ) for two-level model assimilation experiments using relaxation-to-prior perturbation (RTPP) inflation (blue), relaxation-to-prior standard deviation (RTPS) inflation (red), and simple constant covariance inflation (black). The solid lines denote ensemble mean error and the dashed lines denote ensemble spread. The values of the inflation parameter for RTPP and RTPS are given on the lower x-axis, while the values of the constant covariance inflation parameter are given on the upper x-axis. All experiments used covariance localization that tapers covariances to zero 3000 km away from observations locations, and were run for 1000 assimilation steps, after an initial spinup period of 50 days. The definition of the total energy norm is given in the text. 28

440 4 As in Fig. 2, but instead of ensemble mean background error and spread, the ratio  
441 of background spread to analysis spread (the spread growth factor) is plotted as a  
442 function of the inflation parameter. 29

443 5 The mean analysis error covariance singular vector (AECSV) spectrum for RTPP  
444 analysis ensembles with  $\alpha = 0.75$  (red) and RTPS analysis ensembles with  $\alpha =$   
445 0.82 (blue). The solid lines represent the mean growth rates (left y-axis), and the  
446 dashed lines represent the mean background forecast variance explained (right y-  
447 axis) as a function as AECSV singular value number. 30

448 6 Contours of ensemble mean background error using a combination of multiplica-  
449 tive and additive inflation. The additive inflation is created by drawing samples  
450 from a distribution of actual 12-h forecast model errors. The multiplicative infla-  
451 tion parameter varies along the x-axis, while the additive inflation parameter varies  
452 along the y-axis. The solid red line in Fig. 3 is a cross-section along  $y=0$  in this  
453 plot. See text for details. Filter divergence occurs where no contours are plotted. 31

454 7 As in Fig. 6, but additive inflation is created by drawing samples from the clima-  
455 tological distribution of T31 model 12-h. 32

456 8 As in Fig. 7, but for a “perfect model” experiment in which the T42 model is used  
457 in the data assimilation. 33

458 9 As in Fig. 7, but the ensemble size in the data assimilation is increased from 20 to  
459 200. 34

- 460 10 Total kinetic energy spectra as a function of total wavenumber for the T31 model  
461 (blue) the T42 model (green) and the T31 model with stochastic kinetic energy  
462 backscatter (red). See text for details. For reference, a line representing a -3 power-  
463 law spectrum characteristic of 2D turbulence is shown in black. 35
- 464 11 Contours of ensemble mean background error using a combination of multiplica-  
465 tive inflation and stochastic kinetic energy backscatter (SKEB). The multiplicative  
466 inflation parameter varies along the x-axis, while the amplitude of the SKEB forc-  
467 ing varies along the y-axis. The solid red line in Fig. 3 is a cross-section along  
468  $y=0$  in this plot. See text for details. Filter divergence occurs where no contours  
469 are plotted. 36

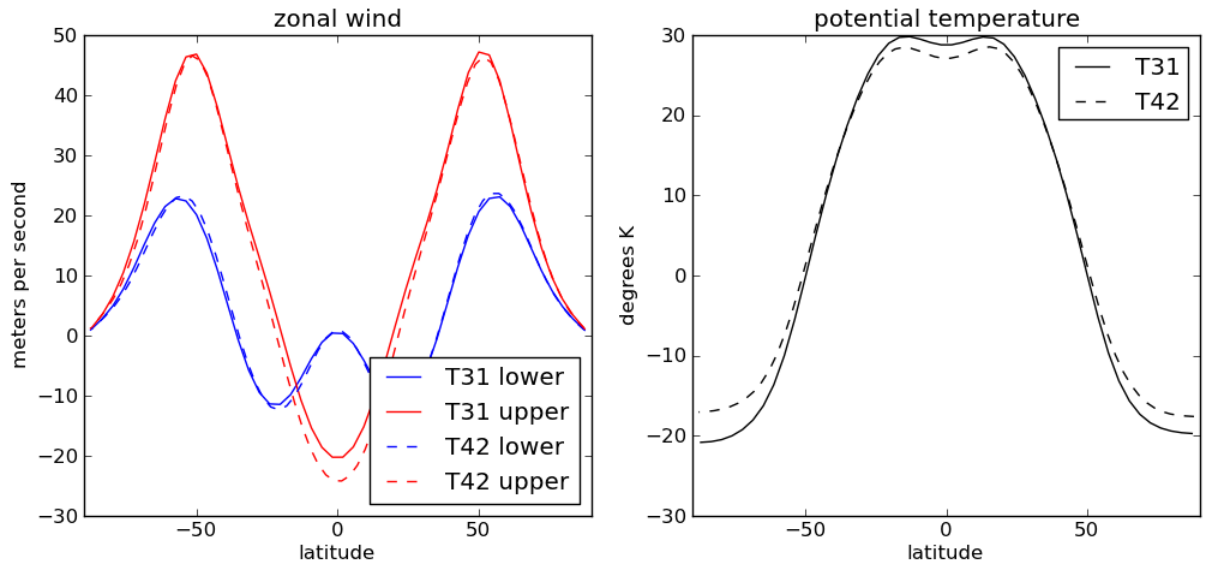


FIG. 1. Zonal and time mean zonal wind (left) and potential temperature (right) for the two-level model run at T31 and T42 resolution.

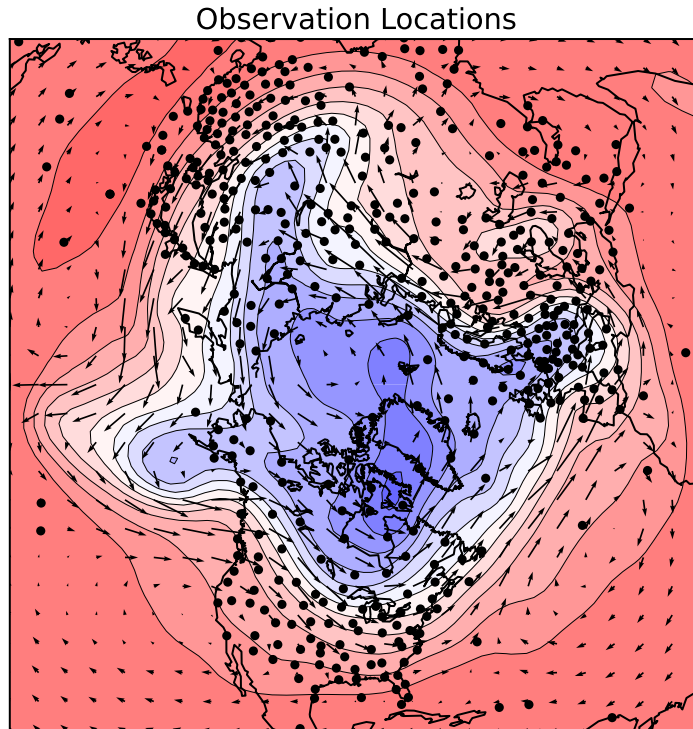


FIG. 2. Observation locations (black dots) superimposed upon a snapshot of barotropic potential temperature (shaded contours every 5K) and upper-level winds (vectors) from the T42 nature run used in the two-level model data assimilation experiments. Continental outlines are shown for reference, even though the model has no orography or land-sea contrast. Data is plotted on a azimuthal equidistant map projection centered on the North Pole.

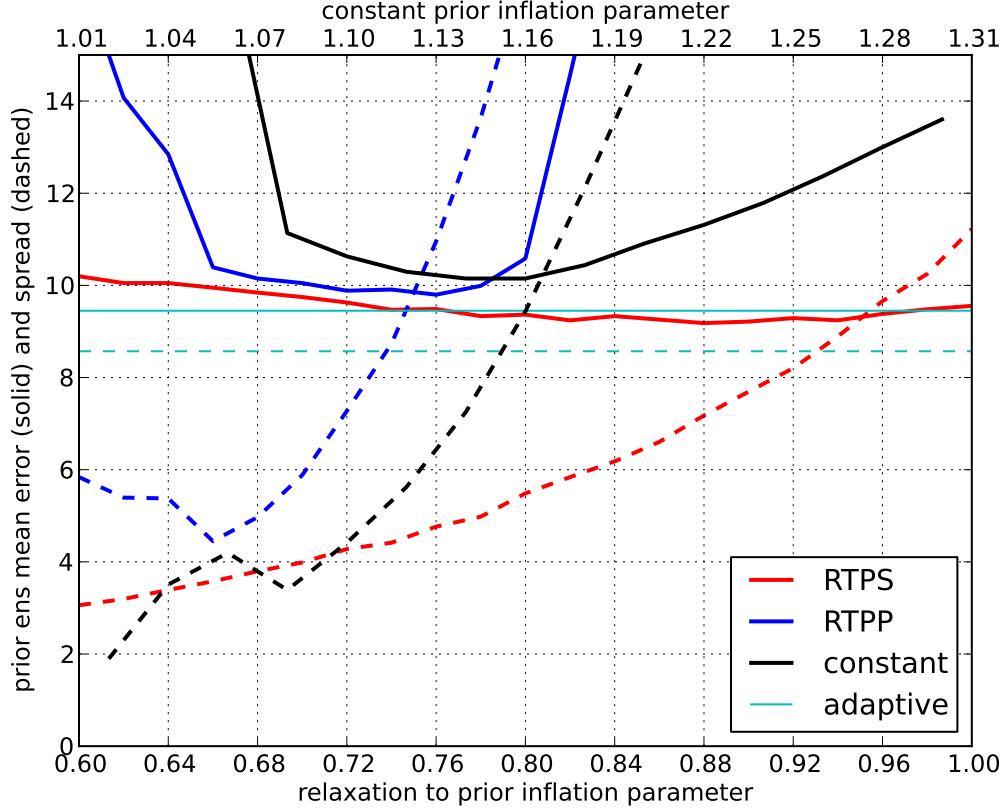


FIG. 3. Ensemble mean background error and spread (in terms of the square-root of total energy norm in  $ms^{-1}$ ) for two-level model assimilation experiments using relaxation-to-prior perturbation (RTPP) inflation (blue), relaxation-to-prior standard deviation (RTPS) inflation (red), and simple constant covariance inflation (black). The solid lines denote ensemble mean error and the dashed lines denote ensemble spread. The values of the inflation parameter for RTPP and RTPS are given on the lower x-axis, while the values of the constant covariance inflation parameter are given on the upper x-axis. All experiments used covariance localization that tapers covariances to zero 3000 km away from observations locations, and were run for 1000 assimilation steps, after an initial spinup period of 50 days. The definition of the total energy norm is given in the text.

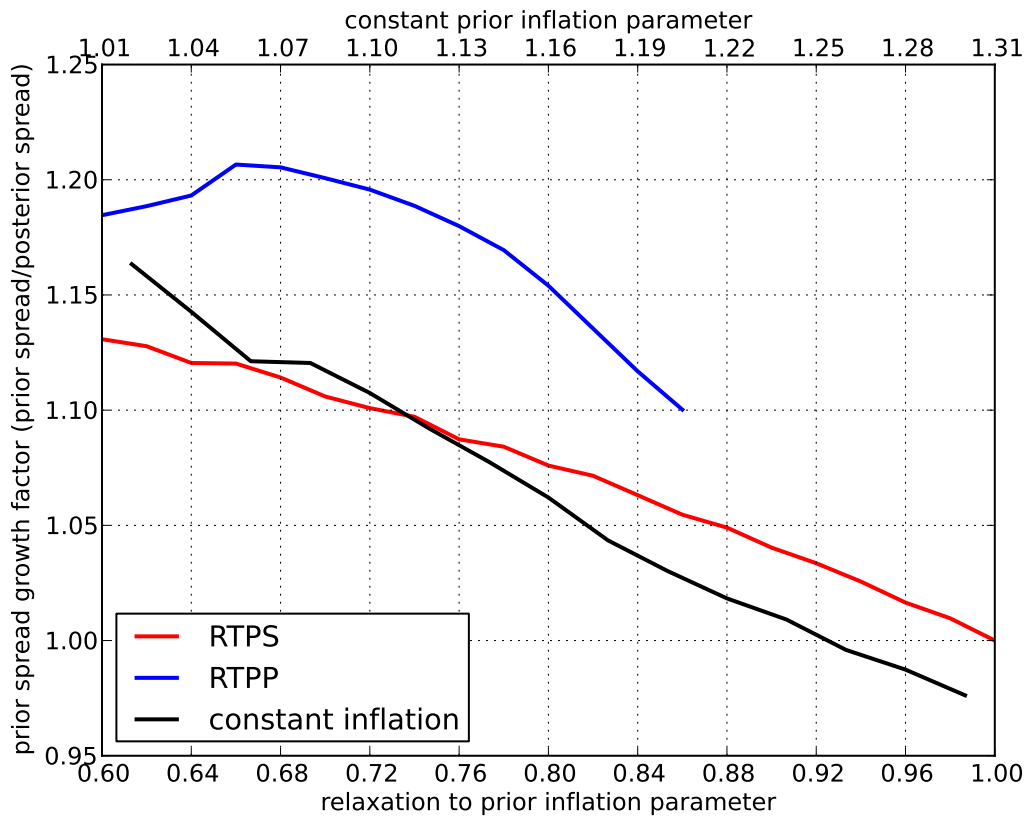


FIG. 4. As in Fig. 2, but instead of ensemble mean background error and spread, the ratio of background spread to analysis spread (the spread growth factor) is plotted as a function of the inflation parameter.

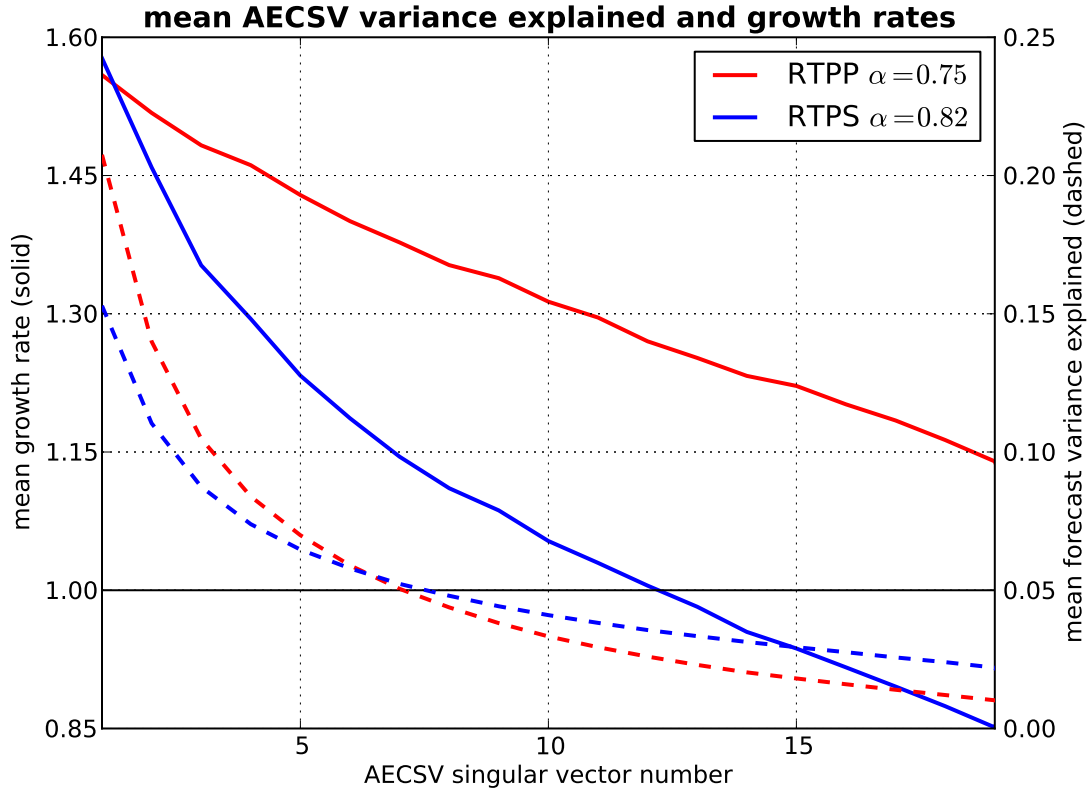


FIG. 5. The mean analysis error covariance singular vector (AECSV) spectrum for RTPP analysis ensembles with  $\alpha = 0.75$  (red) and RTPS analysis ensembles with  $\alpha = 0.82$  (blue). The solid lines represent the mean growth rates (left y-axis), and the dashed lines represent the mean background forecast variance explained (right y-axis) as a function as AECSV singular value number.

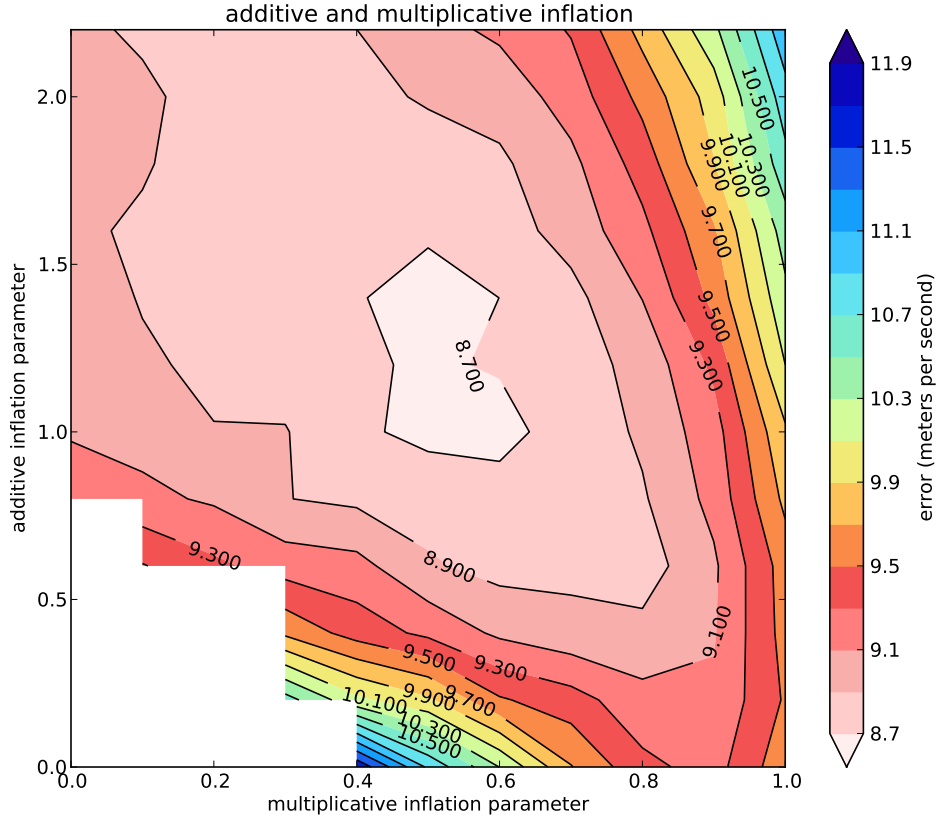


FIG. 6. Contours of ensemble mean background error using a combination of multiplicative and additive inflation. The additive inflation is created by drawing samples from a distribution of actual 12-h forecast model errors. The multiplicative inflation parameter varies along the x-axis, while the additive inflation parameter varies along the y-axis. The solid red line in Fig. 3 is a cross-section along  $y=0$  in this plot. See text for details. Filter divergence occurs where no contours are plotted.

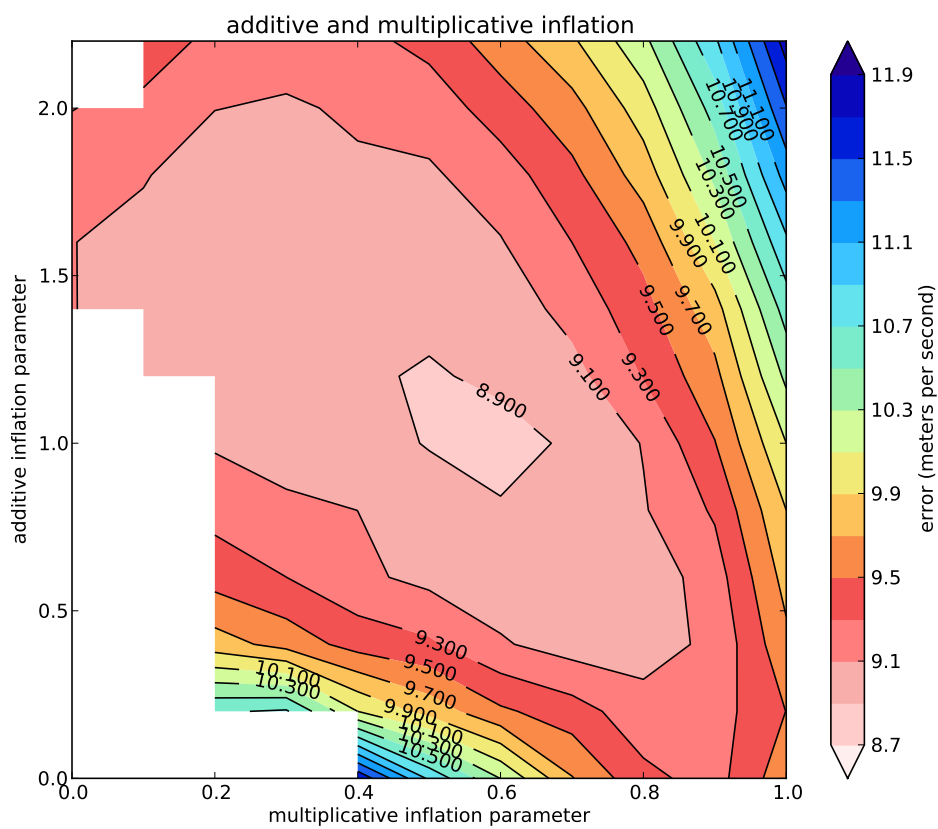


FIG. 7. As in Fig. 6, but additive inflation is created by drawing samples from the climatological distribution of T31 model 12-h.

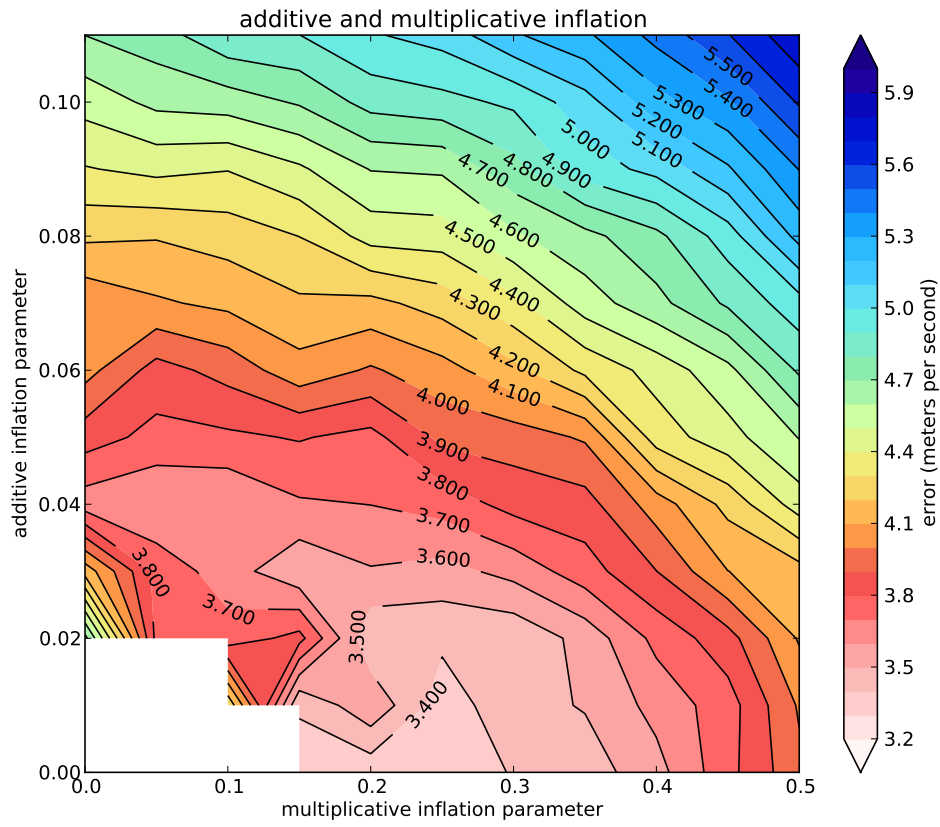


FIG. 8. As in Fig. 7, but for a “perfect model” experiment in which the T42 model is used in the data assimilation.

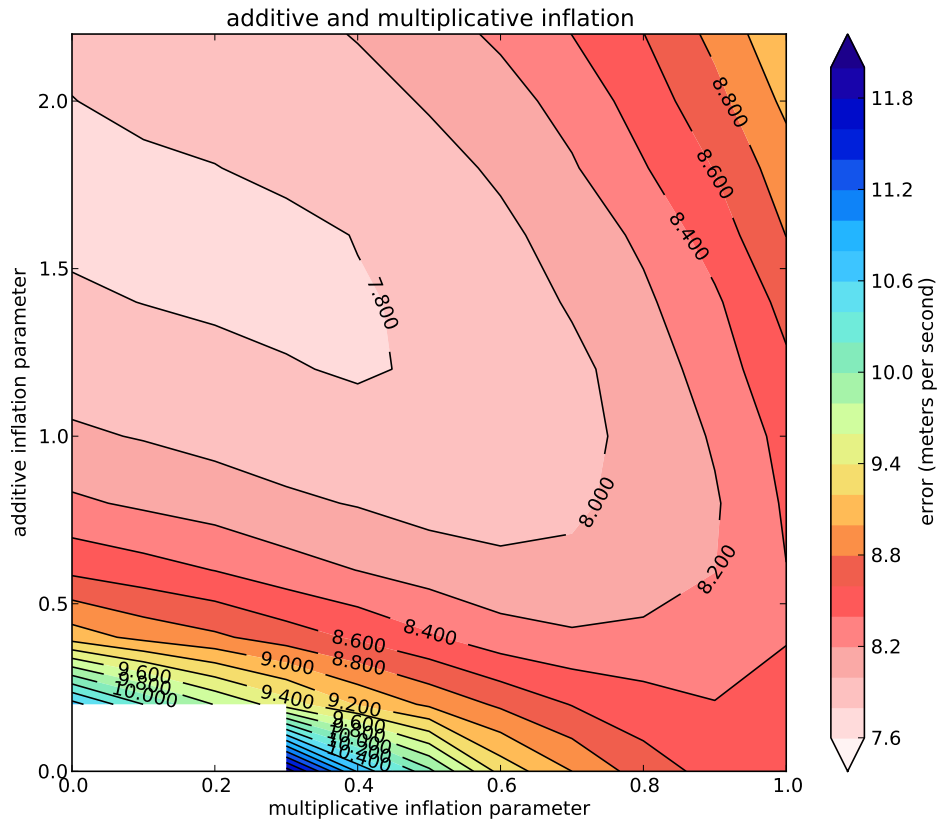


FIG. 9. As in Fig. 7, but the ensemble size in the data assimilation is increased from 20 to 200.

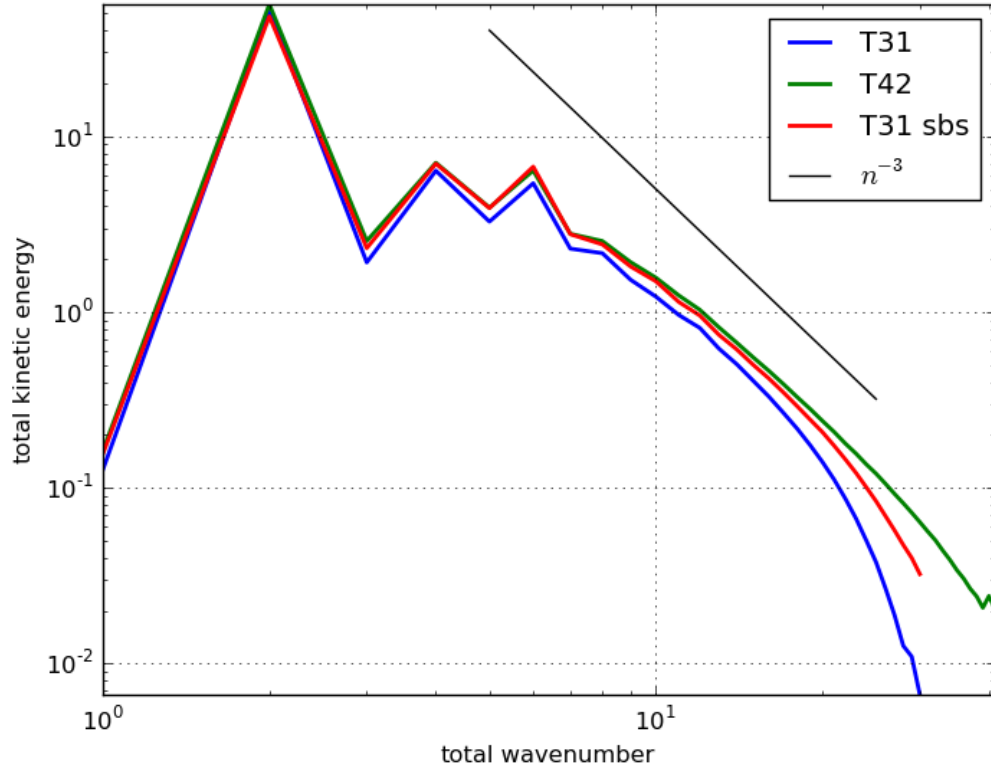


FIG. 10. Total kinetic energy spectra as a function of total wavenumber for the T31 model (blue) the T42 model (green) and the T31 model with stochastic kinetic energy backscatter (red). See text for details. For reference, a line representing a -3 power-law spectrum characteristic of 2D turbulence is shown in black.

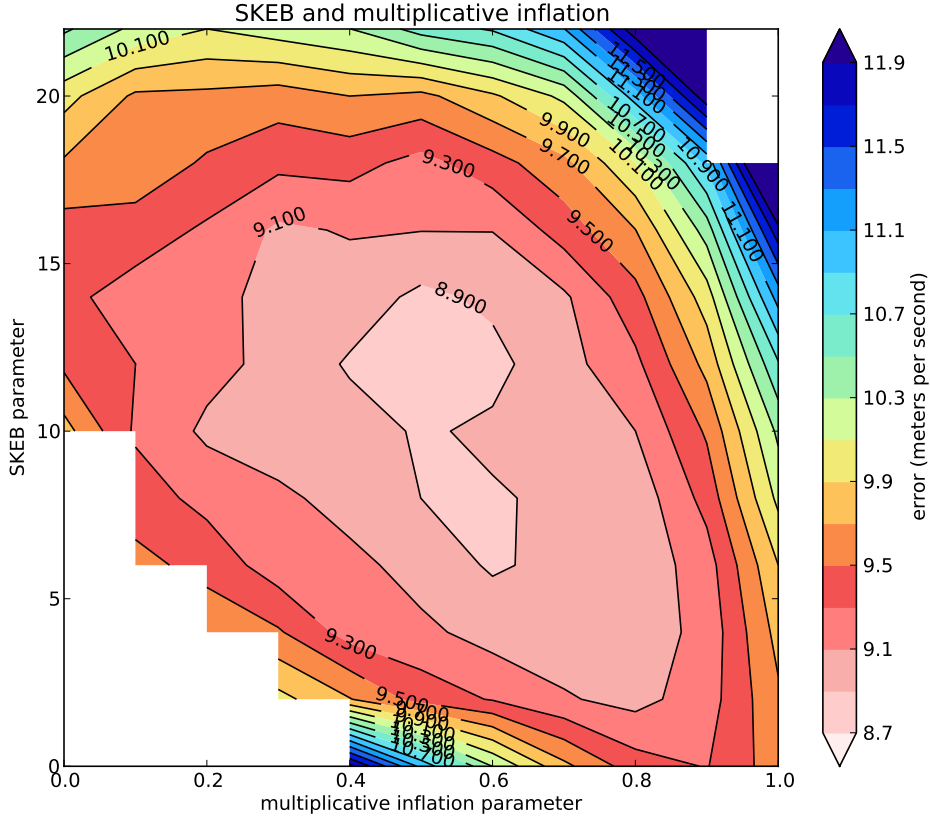


FIG. 11. Contours of ensemble mean background error using a combination of multiplicative inflation and stochastic kinetic energy backscatter (SKEB). The multiplicative inflation parameter varies along the x-axis, while the amplitude of the SKEB forcing varies along the y-axis. The solid red line in Fig. 3 is a cross-section along  $y=0$  in this plot. See text for details. Filter divergence occurs where no contours are plotted.